Marketing Science

Consumer Search and Purchase: An Empirical Investigation of Retargeting Based on Consumer Online Behaviors

Zhenling Jiang, Tat Chan, Hai Che, Youwei Wang

Consumer Search and Purchase: An Empirical Investigation of Retargeting Based on Consumer Online Behaviors

Zhenling Jiang,[a] Tat Chan,[b] Hai Che,[c] Youwei Wang[d]

[a] The Wharton School, University of Pennsylvania, Philadelphia, Pennsylvania 19104; [b] Olin School of Business, Washington University in St. Louis, St. Louis, Missouri 63130; [c] A. Gary Anderson Graduate School of Management, UCR School of Business, University of California, Riverside, Riverside, California 92521; [d] School of Management, Fudan University, Shanghai 200433, China
Contact: zhenling@wharton.upenn.edu, https://orcid.org/0000-0002-4581-0125 (ZJ); chan@wustl.edu (TC); hai.che@ucr.edu,
https://orcid.org/0000-0002-4377-5514 (HC); ywwang@fudan.edu.cn, https://orcid.org/0000-0003-2889-8417 (YW)

Abstract. This paper empirically investigates how marketers can retarget consumers who have searched online but did not purchase, based on their search behaviors. To infer the relationship between search activities and preferences, we estimate a structural search model that characterizes the consumer search process. We propose an estimator similar to the Geweke-Hajivassiliou-Keane estimator to evaluate the likelihood function. The proposed estimator makes recursive draws from truncated distributions that arise because of the observed search and choice behaviors in an optimal sequential search model. The recovered preferences are used to improve retargeting strategies demonstrated through a series of counterfactuals. Results show a substantial heterogeneity in responses to retargeting among consumers who exhibited different search behaviors. By contrast, the heterogeneity among consumers based on other characteristics (e.g., age, gender) is moderate. We consider two counterfactual marketing strategies: sending out coupons redeemed upon purchasing and sending seller recommendations that reveal the offering of recommended sellers. We find that although both strategies help increase the conversion rate, seller recommendations are more effective than coupons, suggesting the importance of providing consumers with the sellers' information for retargeting. We also show that a pricing mechanism such as an auction that makes the seller self-select to participate will improve the effectiveness of retargeting. Finally, online retail platforms can benefit both sellers and consumers by providing sellers with the information on consumers' search behaviors.

Keywords: consumer search • retargeting • GHK estimator • online shopping

1. Introduction

After browsing commercial websites, 95%–98% of individuals leave without taking desirable actions such as purchasing (Berke et al. 2014). Retargeting these individuals has become a common practice in the digital marketing industry. Google Remarketing Campaigns, for example, help marketers track those consumers and reconnect them later by showing display ads when they browse other websites. Retargeted display advertising is found to be effective in bringing consumers back to visit the website (Sahni et al. 2019).

Retargeting display ads are typically shown to all customers who have browsed the website without much differentiation. Customer segmentation or personalization can be a powerful tool in marketing with the right data, such as consumer purchase history (e.g., Fudenberg and Villas-Boas 2006). In online shopping,

advancement in information technology has enabled firms to collect detailed information on individual consumers' online browsing behaviors. Such information has the potential to enable personalization in several ways. First, a firm can prioritize which consumers are more profitable to retarget based on their prior search history. Furthermore, in terms of what to retarget with, it is possible to display a related item that consumers have not seen or an item that consumers have seen but with a lower price (Cristal 2015). Offering the right item or monetary incentives when retargeting can be more effective than the common practice of displaying the last option consumers have browsed. Despite the potential value, little is known from the academic research on how to leverage the search information for retargeting.

This paper intends to fill this gap in the literature. We study how firms can improve the effectiveness of

retargeting by utilizing consumer online search behaviors. The consumer search information can be valuable because firms can infer consumer preferences from their search history, which in turn can be used in retargeting. Results that will be discussed later lend support to this hypothesis. In order to connect consumer search history with their preference, we utilize a consumer search model. The model describes how consumers conduct costly search (in terms of time and effort) to obtain information (Stigler 1961) through browsing different product options. It has been known in the literature that consumer search allows researchers to recover consumer preferences without observing purchases (e.g., Kim et al. 2010). We use this insight to explore effective retargeting strategies, which are different from other retargeting strategies that have been previously studied.

In order to infer preference from consumer search history, we specify and estimate a sequential search model using individual-level data. We contribute to the search model literature by proposing a GHK-type (Geweke et al. 1994, Hajivassiliou and Ruud 1994, Keane 1994) simulation method to estimate sequential search models. The likelihood is the joint probability that describes the observed search and choice behaviors. Unlike the logit-smoothed estimator (e.g., Honka and Chintagunta 2016, Ursu 2018), the proposed estimator evaluates the value of the likelihood function directly without the need to specify scaling factors. It obtains a smooth likelihood function and is more efficient than a crude frequency simulator (e.g., Chen and Yao 2016) where a large number of draws could be rejected. At its core, the GHK estimator is an importance sampling estimator: instead of making draws from the original distribution, it draws from a truncated distribution and applies the importance sampling weight. The area of truncation arises from the set of inequalities characterized by the observed search and choice behaviors. In addition to the simulation method, we also propose a new way to solve the reservation utility with Newton's method in estimating the sequential search model. It is fast and does not introduce approximation errors from the linear interpolation method used in prior literature (e.g., Kim et al. 2010).

This paper studies consumer search behavior and retargeting in the context of an online retail platform where many individual sellers list products to reach final consumers. The empirical data come from Taobao.com, the largest online retail platform in China. The online platform is in a good position to utilize search history to retarget consumers: it collects consumer browsing history and can reach consumers through direct messaging on the platform (through its mobile app or website). This is different from an independent retailer (e.g., Apple.com) that may retarget consumers

who have visited their website before with retargeted display ads (e.g., Sahni et al. 2019). Similar to many other websites, Taobao has the potential to benefit significantly from retargeting efforts because as many as 98.7% of consumers end up not buying after browsing in our data set. If retargeting can convert a small proportion of these individuals, it would have a big boost to its sales and profit.

During the study period, Taobao did not run retargeting. Therefore, it is impossible to rely on historical data to answer the research question.[1] Our approach is to build a consumer search model to establish the structural relationship between consumer preference and search behaviors. After recovering the model primitives, we explore how firms can utilize search history in retargeting through counterfactuals. The counterfactual can provide valuable insights in planning the retargeting and complement other methods such as randomized field experiment. First, it helps narrow down important behavioral dimensions to consider. For example, there are typically many behavioral variables (e.g., How many sellers has a consumer searched? What is the order of the search? What product options have been browsed through during the search?) that can potentially be used. In addition, there are also many possible retargeting strategies to consider (e.g., monetary incentives or product recommendation). Exploring these options directly through methods such as randomized field experiments can suffer from the "curse of dimensionality." Randomized field experiments can also be costly because some of those consumers who are predicted to be profitable will not be retargeted, and vice versa. Another alternative approach is to use statistical methods to predict from the current data how nonpurchasers will respond to retargeting based on the behaviors of purchasers. This approach, however, can suffer from selection issues because the two types of consumers are systematically different. As the structural model has recovered the behavioral primitives, it is able to account for the systematic difference between nonpurchasers and purchasers.

Our empirical search model leverages detailed information on individual-level consumers' search activities and purchase decisions. The model captures the order of search and the conditions of continuing to search or stopping. Furthermore, our model allows for unobserved factors (for researchers) that consumers find at both the search results page and the seller page. We also estimate the value of information, such as consumer reviews and warranty, which consumers (but not researchers) will observe after they click into a seller page. We show how individual-level data on two-stage decisions, clicks and purchases, can identify these unobserved factors among sellers. The model specification is motivated by the empirical

observation that whereas some sellers at Taobao consistently attract many clicks but have low conversion rates, some others are the opposite. Allowing sellers to differentiate in those factors not only helps our model fit with data better but also has important implications on retargeting.

Results show that the cost of searching for information from one seller is ¥26.4 (about US$4.3).[2] Consumers are more likely to search and purchase from sellers who charge a low price and have a high star level. They are also more likely to buy from sellers who live in the same city. Older and male consumers who made a large number of transactions at Taobao before have a higher purchase utility—likewise for those who search during working hours.

We then use the estimation results to investigate the effects of retargeting through a series of counterfactual experiments. We find substantial heterogeneity in the response to retargeting among consumers who exhibited different search behaviors, including which seller links they have clicked, the order of their clicks, and the number of clicks, before being retargeted. By contrast, the heterogeneity among consumers based on other characteristics (e.g., age, gender) is only moderate. These results suggest that search behaviors can carry valuable information for retargeting. We consider two counterfactual marketing strategies. The first is to send out coupons redeemed upon purchasing, and the second is to send seller recommendations via, for example, in-app messaging, which provides details about a seller's actual offering. Results reveal that whereas both strategies help increase profit for Taobao and sellers, the latter is more effective than the former, suggesting the importance of providing consumers with the sellers' information for retargeting campaigns. We also show that Taobao can use a pricing mechanism such as an auction, in which sellers whose offerings can attract consumers will pay to participate, to improve the efficiency of campaigns. Finally, when Taobao provides consumers' search behaviors to sellers, the information will benefit consumers and participating sellers.

Our study is related to three streams of literature. First, it relates to an emerging stream of research on retargeting in marketing. Most of the research in this area measures the effectiveness of retargeted display ads using a randomized field experiment (e.g., Sahni et al. 2019). A few studies have investigated the content of the ads by comparing generic ads and personalized ads, which show the product the consumer most recently browsed over. Personalized retargeting ads are found to be more effective, especially right after the visit (Bleier and Eisenbeiss 2015). Lambrecht and Tucker (2013) find that the effect of personalized ads depends on where consumers are in their product search process—they are more effective for consumers

who are advanced in the process but not for those during early search. Our paper contributes to the retargeting literature in two ways. First, we consider personalized content beyond just the most recently browsed over product. We find that recommending an option that the consumer has not searched can be an effective retargeting strategy. Second, we explore the effectiveness of retargeting among different consumers and find that search history is highly predictive of retargeting effectiveness, more so than other observed characteristics such as age and gender. We study the question with counterfactual results from a structural search model, because the e-commerce website is not currently running retargeting. It will be beneficial to validate the findings using randomized field experiments. We expect that the typical potential negative reaction associated with personalized ads could apply, such as privacy concerns (Goldfarb and Tucker 2011a, 2011b).

Second, the paper is related to the recent empirical literature that estimates consumer search models using individual-level data (e.g., Honka 2014, Koulayev 2014, Chan and Park 2015, and Honka and Chintagunta 2016). This paper makes a methodological contribution to the empirical search literature by proposing a GHK estimator, which obtains a smooth likelihood function without the need for smoothing factors and is more efficient than a crude frequency simulator. Furthermore, for researchers, our model also allows for unobserved factors at both the search results page and the seller page levels, which will impact the researchers' search and purchase decisions. Previous studies have made different model assumptions. Honka (2014), for example, assumes consumers know all firm attributes except prices prior to search, which are observed in her data. Kim et al. (2010), Chen and Yao (2016), and Ursu (2018) allow for independent and identically distributed (i.i.d.) unobserved factors that consumers will find after the search, but there are no unobserved factors before the search. Allowing sellers to differentiate in those factors has important implications on the effectiveness of retargeting.

Finally, our paper is related to the literature on behavioral targeting that utilizes the consumer purchase history. Theoretical research has widely studied the strategic impacts when firms price discriminate consumers with different purchase histories (e.g., Shaffer and Zhang 1995, Villas-Boas 1999, Fudenberg and Tirole 2000, Chen et al. 2001, Villas-Boas 2004, Acquisti and Varian 2005, Pazgal and Soberman 2008, and Shin and Sudhir 2010; for a detailed discussion and review, see Fudenberg and Villas-Boas 2006). Empirical research has used customer-level transaction data to study how firms can estimate individual-specific price elasticity and how to tailor individual-specific prices using targeted coupons (e.g., Elrod and Winer 1982 and Rossi et al. 1996). Cosguner et al. (2016)

find that firms can better tune their targeting strategy given a purchase history because a customer's price elasticity for the most recently purchased brand will decrease at her next purchase occasion as a result of the switching cost. Our paper differs from this stream of literature by investigating how firms can use the information of consumers' online search activities, instead of past purchases, to retarget individuals who have searched but not purchased. Similar to previous studies, we show that such information can help firms better target consumers.

The rest of this paper is organized as follows. We introduce the data and present summary statistics and reduced-form data patterns in Section 2. We describe the sequential search model in Section 3, and we discuss the estimation method in Section 4. Section 5 presents the estimation results. We discuss results from counterfactual retargeting policies in Section 6. Finally, Section 7 concludes.

2. Data Description

The data set used in this study comes from China's largest online retail platform, Taobao.com, which is owned by Alibaba. Individual sellers list products on the platform to sell to end consumers. We focus on the iPad mini 16G WiFi, which has the largest number of searches and purchases in our data. On the Taobao platform, the product is offered by many sellers, who can charge different prices and offer different accessories, warranties, and delivery options. Because a large number of sellers attract only a few clicks and even fewer purchases, we narrow down the focus to the top 20 sellers in terms of the total number of clicks of the product. Our sample of analysis consists of 104,189 consumers who have searched the product among these 20 sellers. All consumers have clicked at least once. The sample of analysis is from March 10 to March 31 in 2013, altogether 22 days.[3]

The data contain the individual-level consumer search process, including the order of search, as well as the purchase decision. We start by describing the consumer search process. A consumer who is interested in buying an iPad submits a query (e.g., iPad mini), and she will be taken to a search results page, which displays a list of links of sellers offering the product. The consumer obtains some information from the product search page. To learn more about a specific option, the consumer clicks on the link and goes to a seller-specific page in a new tab, which contains additional information about the product. Each click on a link represents a seller search. After a search, the customer may decide to either make a purchase, continue the search by clicking another seller link on the search results page, or abandon the search without purchase.

About 11% of consumers have searched at least two options. The intensity of search is positively correlated with the purchase rate. Table 1 shows that 0.9% of consumers who searched only one seller make a purchase. The purchase rate increases to 2.7% for those who searched two sellers and 9.3% for consumers with at least three searches. On average, only 1.26% of consumers ended up making a purchase, implying that 98.7% of consumers who have searched without buying will become the focus of the retargeting effort.

Next we describe what the consumers observe on each search stage and what we as researchers observe. On the search results page, the consumers observe the price, the seller star level, and the geographic location of the seller, which researchers also observe. The star level is calculated from a number of successful transactions a seller has done on the platform. Table 2 reports clicks and purchases of the sellers based on their characteristics. We define the click rate for a seller as the percentage of consumers who have clicked his link. Column (2) reports the average click rate for sellers in each category. We define the conversion rate for a seller as the percentage of consumers who make a purchase conditional on clicking the seller link. Column (3) reports the average conversion rate for sellers in each category. The table shows that consumers are more likely to click and purchase from low-price sellers, sellers with a high star level, and sellers who are from the same geographic location.

On the search results page, consumers observe additional attributes that researchers do not, including a product photo, the title description of the product, and its ranking on the page. After making a click, consumers get to the seller page and obtain further information that is not available to researchers, such as a detailed description of the product, user reviews, and warranty and shipping options. We estimate the value of the unobserved information separately on the search results page and seller page for each seller. The model in Section 3 offers more details.

Across individual sellers, the clicks and purchases vary significantly. We define the click rate for a seller as the percentage of consumers who have clicked his link. Figure 1 shows the click rate and conversion rate of each seller. Although some sellers consistently attract many clicks but are unable to convert them into

Table 1. Number of Searches and Purchase Rate

Number of searches	% of consumers	Purchase rate (%)
1	88.9	0.9
2	8.9	2.7
3	1.6	7.7
4	0.4	14.3
5 or more	0.1	12.7

Table 2. Clicks and Purchases Based on Seller Characteristics

Seller characteristics	Proportion of sellers (%)	Click rate (%)	Conversion rate (%)
Price			
Low (average ¥2,106)	50	7.61	1.84
High (average ¥2,312)	50	3.80	1.60
Seller star level	Proportion of sellers (%)	Click rate (%)	Conversion rate (%)
Low (average 9.1)	50	1.89	1.65
High (average 13.4)	50	8.82	1.78
Geographic location	Proportion of consumers (%)	Click rate (%)	Conversion rate (%)
Same as consumer's	2.8	9.35	3.69
Different from consumer's	25.5	5.40	1.94
Consumer's missing	71.7	5.66	1.57

purchases, some others are just the opposite. For example, the number of clicks on seller 1's link is more than five times higher than that on seller 4's, but his conversion rate is just about one-third of seller 4's. Seller 14 has a very low click rate, but his conversion rate is the highest. The large variation in click rates across sellers can be driven by the seller ranking, which can affect consumers' search behaviors (Ursu 2018). Figure 2 plots the click rate per day for each seller across the 22 days. The variation in click rate over time is small for the majority of sellers, and it is dwarfed by the large difference in click rates across sellers. The small variation over time implies that the seller ranking is likely stable over the 22-day sample period.[4]

Besides seller attributes, we also observe some consumer characteristics. Demographic information, including age and gender, is self-reported by about 60% of consumers. The data contain the number of past transactions the consumers have made at Taobao. We also observe the timing of the search—whether the search is conducted during working hours and whether it is during weekend. Table 3 reports the summary statistics of consumer characteristics and search and purchase behaviors. Column (3) shows the average number of searches, defined as the number of clicks on seller links, and

column (4) shows the purchase rate, defined as the percentage of consumers who make a purchase. The average number of searches is quite close among different consumer groups, but the conversion rates vary substantially. Older male consumers with a large number of previous transactions at Taobao have a higher purchase rate. Furthermore, consumers who search during weekdays and during working hours are more likely to make purchases.

Finally, we examine the relationship between the probability of purchase and the order of search, which motivates us to use a sequential search model to describe the consumer search behavior. Figure 3 shows that there is a U-shaped relationship between the order of search and the conversion rate, which is defined as the proportion of consumers who purchase from a seller conditional on click. In other words, consumers are more likely to buy from the first or the last seller they have searched than the ones in the middle of the search process. The U-shaped relationship is consistent with an optimal sequential search model (Weitzman 1979).[5] Consumers start searching from the option with the highest reservation value, which is the one with the highest expected value, if the search cost is the same for each option. Therefore, the conversion rate for the first clicked option should be high because it is the one that looks

Figure 1. Click Rates and Conversion Rates Across Sellers
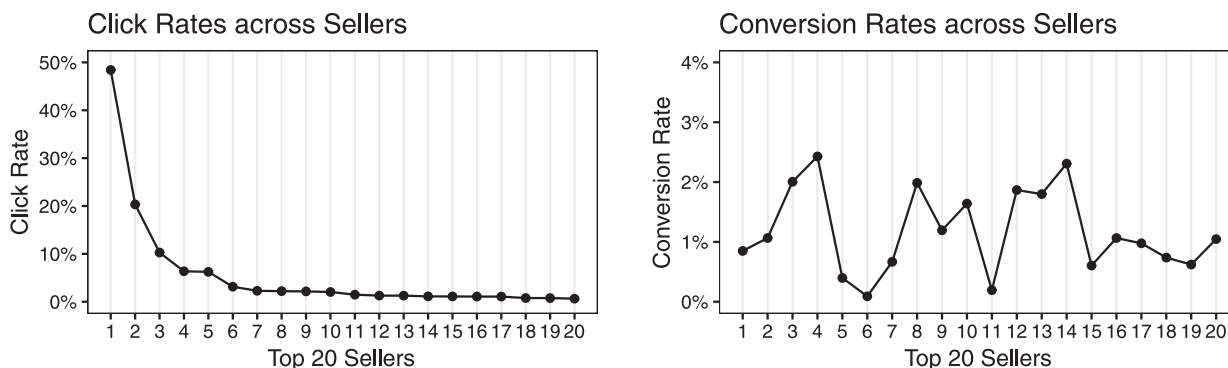
Figure 2. (Color online) Click Rate for Each Seller over Time
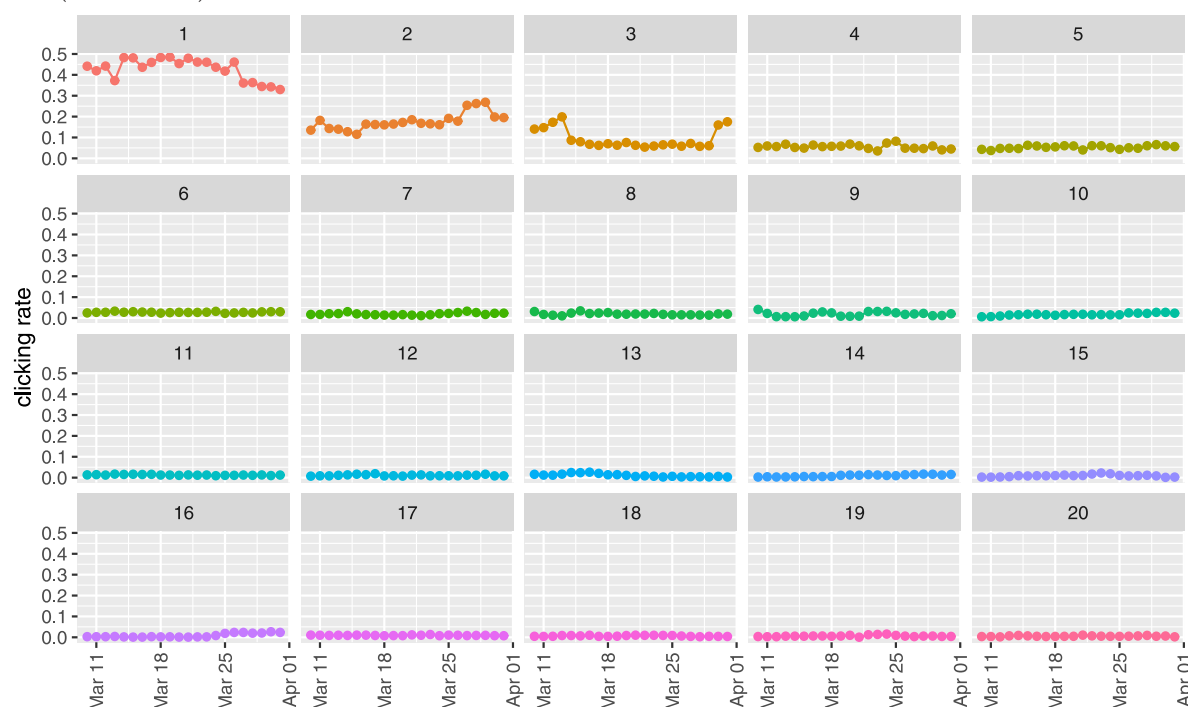
the best ex ante. For the last clicked option, the conversion rate should also be high if consumers search sequentially. This is because consumers are more likely to stop the search after finding a good option.

3. Model

To study how sellers at Taobao can efficiently retarget consumers who have searched but not purchased, we infer their preferences from search behaviors prior to retargeting. We build a model based on Weitzman's optimal sequential search model (Weitzman 1979), which applies to our empirical context with differentiated options to search from. We explicitly model the consumer decisions of which seller link to click and, after clicking a seller page, whether to continue searching or stop and whether to purchase from the seller.

Table 3. Clicks and Purchases Based on Consumer Characteristics

Consumer characteristics	Number of consumers	% of consumers	Average number of searches	Purchase rate (%)
Past number of transactions				
0–30	23,714	22.6	1.12	0.86
31–150	41,631	39.7	1.13	1.08
151 or above	39,474	37.7	1.16	1.67
Gender				
Male	37,930	36.2	1.16	1.51
Female	23,882	22.8	1.14	1.27
Missing	43,007	41.0	1.12	1.02
Age				
18–25	26,179	25.0	1.15	1.13
26–30	17,333	16.5	1.15	1.60
31 or above	18,300	17.5	1.15	1.67
Missing	43,007	41.0	1.12	1.02
Time of visit				
Working hours	63,213	60.3	1.15	1.50
Nonworking hours	41,606	39.7	1.13	0.88
Day of visit				
Weekday	71,524	68.2	1.14	1.36
Weekend	33,295	31.8	1.14	1.04

Figure 3. Order of Search and Conversion Rate
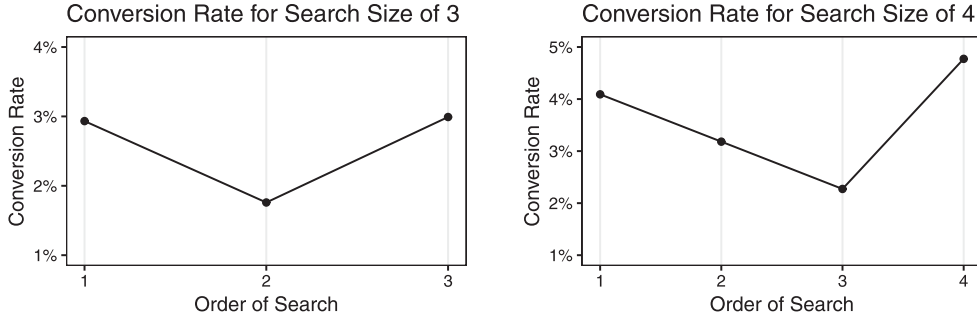
3.1. The Purchase Utility

The value from a seller's offering is captured by a purchase utility function. The utility consumer i obtains from buying from seller j has two components:

$$u_{ij} = V_{ij} + \delta_{ij}. \tag{1}$$

The first component, V_{ij}, is observed by the consumer at the search results page, whereas the second component, δ_{ij}, is revealed after she clicks into the seller's page. The value of the no-purchase option, u_{i0}, is normalized to 0.

The value observed at the search results page, V_{ij}, has several components. First, as discussed in Section 2, the consumer observes some product attributes that researchers also observe, such as the price, the seller's star level, and the location of the seller. In addition, she also obtains information that researchers do not observe, such as the product ranking on the page and whether the product photo looks appealing. Let V_{ij} be

$$V_{ij} = \beta_0 + X_{ij}\beta + \xi_j + e_{ij}. \tag{2}$$

Here, X_{ij} captures the variables that consumers and researchers both observe on the search results page, and β is a vector of parameters. The attributes unobserved to researchers (but observed by consumers) can be further decomposed into two parts. The first part, ξ_j, represents the latent factors that are common to consumers for seller j, such as whether an option is ranked high on the search results page. Beyond what the observed attributes can explain, a seller with a higher ξ value is likely to receive more clicks. This is different for the second part, e_{ij}, which represents an idiosyncratic shock that is i.i.d. across consumers and sellers. Note that ξ_j includes the impact from the seller's ranking on the page. We assume that the consumer perceives the ranking as a signal of the value of the seller's offering (e.g., Bentley et al. 2018), and we draw inference that a top-ranked option is appealing to most buyers and carries a higher perceived value.[6]

After the consumer clicks into the seller page, the actual purchase utility is fully realized. The consumer observes the new information, δ_{ij}, on the seller's page.

Similar to the preclick stage, δ_{ij} can be further decomposed into two parts:

$$\delta_{ij} = \omega_j + \epsilon_{ij}. \tag{3}$$

The first part, ω_j, represents the value of the information on the seller's page that is common across consumers. For example, if a seller has positive reviews or an easy return policy, such information is valued by all consumers and is captured by a high ω_j for the seller. Although we as researchers do not observe the information on the seller's page, we estimate each ω_j. The ω_j term allows the information on the seller's page to be systematically higher for some sellers than others. This is consistent with the different conversion rates conditional on click across sellers (Figure 1). The second part, ϵ_{ij}, is an idiosyncratic shock that is i.i.d. across consumers and sellers. Because both e_{ij} are ϵ_{ij} are assumed to be i.i.d., they are independent from each other. It is conceivable that these two error terms may be correlated. The findings in the paper are robust to a potential correlation between e_{ij} and ϵ_{ij} (see Online Appendix 1 for details).

Consumers search to find out δ_{ij}. We assume that consumers know the distribution of $\epsilon_{ij} \sim N(0, \sigma_\epsilon^2)$, and they have rational expectation of the distribution of ω_j. In model estimation, we estimate ω_j for each seller with the normalization that the average ω_j across sellers is 0. We assume that the $\omega's$ for all sellers are drawn from a normal distribution $N(0, \sigma_\omega^2)$ and calculate the dispersion of ω_j as $\sigma_\omega^2 = \sum_{j=1}^{K} \frac{\omega_j^2}{K-1}$, where K is the total number of sellers.[7] Therefore, before a click, consumers perceive δ_{ij} to be normally distributed with mean 0 and variance $\sigma_\delta^2 = \sigma_\omega^2 + \sigma_\epsilon^2$. It satisfies the rational expectation assumption because the estimated ω_j for each seller indeed has the variance of σ_ω^2.

3.2. The Optimal Search Strategy

Let c_i be the individual-specific marginal cost of searching an additional seller page. The cost captures the time and effort the consumer spends on learning the information from the seller page. To ensure that it

is positive, we assume that c_i across consumers follows a log-normal distribution. That is,

$$\log(c_i) \sim N(\bar{c}, \sigma_c^2). \tag{4}$$

Following Weitzman (1979), to decide which seller page to click, the consumer first forms a reservation utility z_{ij} for every seller on the search results page and then chooses the seller with the highest z_{ij}. The reservation utility is determined by the following:

$$c_i = \int_{z_{ij}}^{\infty} (V_{ij} + \delta_{ij} - z_{ij}) f(\delta_{ij}) \, d\delta_{ij}. \tag{5}$$

The optimal search strategy has three components: the selection rule, the stopping rule, and the choice rule. The consumer will order sellers based on the reservation utility and start searching the option with the highest reservation utility. She will stop searching if the highest utility searched so far or the outside option (normalized to 0) is greater than the maximum of the reservation utilities for unsearched options; otherwise, she will continue to search. After the consumer stops searching, she will choose the option with the highest utility among the set of sellers she has searched and the no-purchase option.

If the consumer decides to terminate the search without purchase, she becomes the potential target consumer for retargeting. The optimal search rules allow us to infer the preferences of the retargeting consumers from their search activities. Specifically, the order of the search sequence and the set of sellers a consumer chooses to search through provide information on the expected utilities V_{ij} for sellers the consumer chooses to search or not to search. The fact that she chooses not to purchase provides further information on her actual utility u_{ij} for sellers for whom she has searched. Finally, the purchase decisions of other consumers help project the value u_{ij} for sellers who have not been searched for. With this information, we can predict how retargeting efforts are able to attract the consumer back to search again and make a purchase. We will further discuss the details in the results section.

4. Estimation
4.1. Calculate Reservation Utility
We first describe how to calculate reservation utility z_{ij} for consumer i with search cost c_i. Let σ_δ be the standard deviation of the portion of the utility that is revealed after search. The normalized reservation utility is $\zeta_i = \frac{z_{ij} - V_{ij}}{\sigma_\delta}$, and the normalized search cost is $x_i = \frac{c_i}{\sigma_\delta}$. Note that with the additive utility specification, the normalized reservation utility ζ_i is only a function of the normalized search cost, and it does not depend on the portion of the utility that is observed

prior to search V_{ij} (see Choi et al. 2018). From Equation (5), the relationship between the normalized search cost and normalized reservation utility can be written as follows (see Kim et al. 2010):

$$x_i = \mathscr{B}(\zeta_i) = (1 - \Phi(\zeta_i))\left(\frac{\phi(\zeta_i)}{1 - \Phi(\zeta_i)} - \zeta_i\right). \tag{6}$$

Kim et al. (2010) show that $\mathscr{B}$ is a monotonic function, and there exists a unique reservation utility that solves $\zeta_i = \mathscr{B}^{-1}(x_i)$. During estimation, one needs to calculate the corresponding reservation utility given some search cost. The reservation utility has no closed-form solution and is therefore difficult to calculate directly. Prior literature solves this problem by pre-computing a table of search costs and reservation utilities and uses linear interpolation to search for ζ_i (e.g., Kim et al. 2010 and Chen and Yao 2016).

We show that the reservation utility can be calculated using Newton's method, without the approximation error from the linear interpolation method. Solving for ζ_i for a given x_i in Equation (6) is equivalent to finding the solution for the function $g(\zeta_i) = (1 - \Phi(\zeta_i))(\frac{\phi(\zeta_i)}{1 - \Phi(\zeta_i)} - \zeta_i) - x_i = 0$. Newton's method uses numerical analysis to find successively better approximations to the root of a function. The algorithm starts with an initial guess and iteratively finds the next guess as $\zeta_{i,k+1} = \zeta_{ik} - \frac{g(\zeta_{ik})}{g'(\zeta_{ik})}$. The iteration process stops when $\zeta_{i,k+1}$ and ζ_{ik} are very close.[8] The function $g(\zeta_i)$ is continuous and differentiable. After plugging in $g'(\zeta_i)$ and rearranging terms, the next guess in the iteration can be simplified to

$$\zeta_{i,k+1} = \frac{\phi(\zeta_{i,k}) - x}{1 - \Phi(\zeta_{i,k})}.$$

See Appendix Section A.1 for a detailed derivation. The convergence is very fast, requiring six or seven iterations on average in our empirical application. With a slight abuse of notation, we use $\zeta(x_i)$ to denote the solution of normalized reservation utility ζ_i for a given normalized search cost x_i. The reservation utility is then $z_{ij} = V_{ij} + \sigma_\delta \cdot \zeta(x_i)$.

Beyond calculating reservation utility, we expect that Newton's method can be useful in other scenarios where the direct solution does not exist but it is easy to calculate the first-order derivative. It can be a useful alternative approach to linear interpolation in order to reduce approximation error.

4.2. Simulated Maximum Likelihood
In the model, the parameter set includes β_0, β, ξ_j (Equation (2)), ω_j (Equation (3)), and $\bar{c}$ and σ_c^2 (Equation (4)). We estimate the model using the maximum likelihood approach.

Let J be the total number of searches the consumer makes, and let j represent the order of searches, such that $j = 1$ for the first clicked option and $j = J$ for the last clicked option. The total number of options to choose from is K. After the consumer clicks J options, we denote the chosen option as $j^\star$ (including the no-purchase option 0). The probability of observing a certain outcome can be characterized by the joint probability of the components of optimal sequential search:

• *Order of search*: The reservation utility of the first clicked option z_{i1} is higher than that of the second clicked option z_{i2}, and so on. Furthermore, the reservation utility of all the clicked options is higher than the maximum from the unsearched set:

$$z_{i1} > z_{i2} > \cdots > z_{iJ} > \max_{j' \notin (1,\dots,J)} \left(z_{ij'}\right). \quad (7)$$

• *Continue to search or stop searching*: The consumer will continue to search when the maximum utility from the searched options and the no-purchase option is smaller than the maximum reservation utility from unsearched options (when $J > 1$):

$$\max(u_{i0}, u_{i1}) < z_{i2}, \dots, \max(u_{i0}, u_{i1}, \dots, u_{i,J-1}) < z_{iJ}. \quad (8)$$

Otherwise, the consumer will stop searching:

$$\max(u_{i0}, u_{i1}, \dots, u_{iJ}) > \max_{j' \notin (1,\dots,J)} \left(z_{ij'}\right). \quad (9)$$

• *Choice*: The consumer will choose the option with the highest utility from all searched options and the no-purchase option:

$$u_{ij^\star} \geq \max(u_{i0}, u_{i1}, \dots, u_{iJ}). \quad (10)$$

The likelihood of observing a consumer's search and purchase behaviors is the joint probability that Equations (7)–(10) hold true:

$$
\begin{aligned}
L = \prod_{i=1}^{N} \int \Big\{ & z_{i1} > z_{i2} > \dots > z_{iJ} > \max_{j' \notin (1,\dots,J)} \left(z_{ij'}\right), \\
& \max(u_{i0}, u_{i1}) < z_{i2}, \dots, \max(u_{i0}, u_{i1}, \dots, u_{i,J-1}) \\
& < z_{iJ}, \max(u_{i0}, u_{i1}, \dots, u_{iJ}) > \max_{j' \notin (1,\dots,J)}(z_{ij'}), \\
& u_{ij^\star} \geq \max(u_{i0}, u_{i1}, \dots, u_{iJ}) \Big\} dF(c_i) \prod_{j=1}^{K} dF(e_{ij}) dF(\epsilon_{ij}).
\end{aligned}
$$
$$(11)$$

where $F(c_i)$ is the distribution function of c_i, and $F(e_{ij})$ and $F(\epsilon_{ij})$ are the distribution functions of e_{ij} and ϵ_{ij}, respectively.

4.3. A GHK Simulator
We write out the likelihood function in terms of the observed variables and error terms that need to be integrated out. Recall that the reservation utility is $z_{ij} = V_{ij} + \sigma \cdot \zeta(x_i)$, where x_i is the normalized reservation utility $x_i = \frac{c_i}{\sigma}$. Furthermore, V_{ij} can be decomposed into two parts, $V_{ij} = \bar{V}_{ij} + e_{ij}$, where e_{ij} is an i.i.d. random shock that consumers observe before search (see Equation (2)).

For estimation purposes, we separate consumers into three types based on their purchase behavior: (1) nonpurchasers, (2) those who purchase the last clicked option, and (3) those who purchase an option that is not the last clicked. This is because the likelihood function can be simplified based on the type of purchase behavior. We illustrate the proposed GHK simulator using an example of cases where a consumer does not make a purchase. The logic is the same for the other two types with different simplification based on purchase behavior. For nonpurchasers, the conditions in Equations (7)–(10) can be simplified as follows.

4.3.1. Nonpurchasers ($j^\star = 0$). We have the following conditions:
• The order condition (Equation (7)):

$$\bar{V}_1 + e_1 > \bar{V}_2 + e_2 > \dots > \bar{V}_J + e_J > \max_{j' \notin (1,\dots,J)} \left(\bar{V}_{j'} + e_{j'}\right). \quad (12)$$

• The continue to search condition when $J > 1$ Equation ((8)):

$$0 < \bar{V}_J + e_J + \sigma \cdot \zeta(x). \quad (13)$$

The stop searching condition (Equation (9)):

$$\max_{j' \notin (1,\dots,J)} \left(\bar{V}_{j'} + e_{j'}\right) + \sigma \cdot \zeta(x) < 0. \quad (14)$$

• The choice condition:

$$\max_{j \in (1,\dots,J)} \left(\bar{V}_j + e_j + \omega_j + \epsilon_j\right) < 0. \quad (15)$$

For nonpurchasers, the likelihood is the joint probability that these four conditions all hold true. The likelihood function does not have a closed-form solution and has to be evaluated via simulations. The likelihood function integrates over high-dimensional random variables, with K e's as well as K ϵ's and c_i.

A straightforward way to evaluate the likelihood function is to use the crude frequency simulator. One can draw S sets of random variables unconditionally from their respective distributions, $e_k^{(s)}, \epsilon_k^{(s)}, c^{(s)}$ $k = 1, \dots, K$, where K is the number of available options. The crude frequency simulator is the proportion that the four conditions are satisfied. In practice, however, using the crude frequency simulator can be very inefficient because a large number of draws would be rejected. For example, the e's draws need to satisfy the observed

order of clicks (Equation (12)) and the continue to search and stop searching conditions (Equations (13) and (14)). Given a finite number of draws, the estimated probability can be 0 or imprecisely estimated. This is especially the case for low-probability conditions (e.g., unpopular search sequences).

Prior literature solves this problem by using a logit-smoothed estimator (Honka and Chintagunta 2016, Ursu 2018). Instead of counting the proportion that all conditions are satisfied, they convert the inequalities into expressions that are positive if the conditions are satisfied and negative otherwise, and they apply a logit function to obtain smoothed probabilities. When using a logit-smoothed estimator, each expression needs to be weighted by a scaling factor. Choosing the "right" scale factors can be challenging. It is especially the case for sequential search models where the scaling factors will implicitly determine the relative weight of the conditions (e.g., order versus continue to search conditions) when constructing the overall simulated likelihood.[9]

We propose a GHK simulator to evaluate the likelihood function. The GHK simulator (Geweke et al. 1994, Hajivassiliou and Ruud 1994, Keane 1994) has been used to draw truncated multivariate variables with correlation, such as in the multinomial probit model. In the probit model, the error terms need to satisfy a set of inequalities that are imposed from the observed choice condition. The GHK simulator is an importance sampling estimator: instead of making draws from the original distribution (e.g., $\phi(x)$), we recursively draw the error terms from a truncated distribution (e.g., $\frac{\phi(x)}{\Phi(\bar{b}) - \Phi(\underline{b})}$) and apply the importance sampling weight (e.g., $\Phi(\bar{b}) - \Phi(\underline{b})$). One can first perform Cholesky decomposition of the variance-covariance matrix and recursively draw each error term, because the truncated region of one error term depends on the draws of the other error terms. In a sequential search model, such inequalities arise from the observed consumer search and choice conditions. From the perspective of the researchers, the distributions of the error terms are truncated, with the truncated region coming from the inequality conditions as defined in Equations (12)–(15). Similar to the case of the multinomial probit model, in the sequential search model, the area of truncation is also recursively determined.

We describe the sampling procedure using the example of nonpurchasers (see Appendix Section A.2 for the sampling procedure for purchasers). Recall that $e's \sim N(0, \sigma_e^2)$ with the probability density function (PDF) denoted by as ϕ^e and cumulative distribution function (CDF) denoted as Φ^e,[10] and $\epsilon's \sim N(0, \sigma_\epsilon^2)$, where σ_ϵ is normalized to 1 (see Kim et al. 2010) with PDF ϕ and CDF Φ.

1. Sampling c is straightforward. We draw $c^{(s)}$ from the log-normal distribution (Equation (4)) and calculate the normalized search cost $x^{(s)} = \frac{c^{(s)}}{\sigma}$.

2. Sample $e_{j'}$'s for the unsearched options, whose area of truncation comes from the stop searching condition (Equation (14)).

    a. Draw $e_{j'}^{(s)}$'s, $\forall j' \notin (1,\ldots,J)$ from their distributions with upper bound $\bar{b} = -\bar{V}_{j'} - \sigma \cdot \zeta(x^{(s)})$.

3. Sample e_J for the last searched option, whose truncated region comes from both the order condition (Equation (12)) and continue to search condition (Equation (13)).[11] The truncated region depends on the $e_{j'}^{(s)}$'s draws.

    a. Draw $e_J^{(s)}$ with lower bound $\underline{b} = \max_{j' \notin (1,\ldots,J)} (-\bar{V}_J - \sigma \cdot \zeta(x^{(s)}), -\bar{V}_J + \bar{V}_{j'} + e_{j'}^{(s)})$.

4. Sample $e_{J-1}, \ldots, e_1$ recursively (if $J > 1$). The draws need to satisfy the order condition (Equation (12)), which depends on the previous draws of $e_j^{(s)}$.

    a. Recursively draw $e_j^{(s)}$ from its distribution with lower bound $\underline{b} = \bar{V}_{j+1} + e_{j+1}^{(s)} - \bar{V}_j$, where $j = J - 1, \ldots, 1$.

5. We then evaluate the likelihood of the choice condition (Equation (15)). The probability that $\epsilon's$ satisfy the inequalities can be directly evaluated:

    a. $P(\epsilon_j < -\bar{V}_j - e_j^{(s)} - \omega_j) = \Phi(-\bar{V}_j - e_j^{(s)} - \omega_j)$, $\forall j \in (1, \ldots, J)$.

6. Finally, the likelihood can be approximated by the average of the probabilities using a set of draws:

$$L^s = \frac{1}{S} \sum_{s=1}^{S} \left\{ \prod_{j' \notin (1,\ldots,J)} \Phi^e\left(-\bar{V}_{j'} - \sigma \cdot \zeta\left(x^{(s)}\right)\right) \cdot \right.$$

$$\left(1 - \Phi^e\left(\max_{j' \notin (1,\ldots,J)} \left(-\bar{V}_J - \sigma \cdot \zeta\left(x^{(s)}\right), -\bar{V}_J + \bar{V}_{j'} + e_{j'}^{(s)}\right)\right)\right) \cdot$$

$$\left. \prod_{j=1}^{J-1} \left(1 - \Phi^e\left(\bar{V}_{j+1} + e_{j+1}^{(s)} - \bar{V}_j\right)\right) \cdot \prod_{j=1}^{J} \Phi\left(-\bar{V}_j - e_j^{(s)} - \omega_j\right) \right\}.$$

The proposed estimator is much more efficient than the crude frequency simulator, and it does not require choosing scaling factors as in the logit-smoothed estimator. Using the GHK sampling method is especially beneficial when the dimensionality of integration is high with truncation, where simple unconditional sampling would lead to many draws being rejected. Furthermore, the simulated likelihood function is smooth and differentiable, which makes the model estimation converge faster.

4.4. Identification
4.4.1. Preclick Utility V_{ij}.
The identification of the parameters in V_{ij} prior to clicking comes from the consumers' clicking and purchase behaviors across sellers. The systematic relationship between X_{ij} and

the click rate identifies β in Equation (2) similar to a choice model. Because Equation (2) includes an intercept, β_0, we normalize the average ξ_j across sellers to 0 (more details are in the paragraph below). Furthermore, the utility of no purchase is normalized to 0. With this normalization, the average conversion rate in data identifies the intercept parameter β_0.

With the preclick seller fixed effects ξ_j in the model, one needs to use variation in X's within seller to identify β. During the sample period, the price rarely changes, and the star level remains mostly the same. Without further assumptions we cannot separate the effects of prices and seller stars from ξ_j on consumers' utility. Our strategy is to rely on the variations of prices and stars across sellers. We use the median criterion to categorize sellers into a high-price group and a low-price group, as well as a high-star group and a low-star group. Assuming that ξ's are not systematically different between the groups, we normalize the average ξ_j in the high- and low-price group to 0. Similarly, the average ξ_j in the high- and low-star group is also normalized to 0. As such, the coefficient for the indicator of the high-price group represents how likely consumers are to click and purchase from a seller with a high price relative to the low-price group. The coefficient for the indicator for the high-star group is similarly interpreted. The estimated ξ_j for a seller represents his ability to attract clicks in comparison with the other sellers in the same group.[12]

For the variance of the idiosyncratic term σ_e^2, conceptually it can be identified from the variation in clicking choices across consumers. If σ_e^2 is small, the differences in the perceived value before clicking across consumers is small, in which case consumers faced with the same list of sellers will follow a very similar search sequence. Otherwise, if σ_e^2 is large, then the search sequence across consumers can be very different. In the estimation, however, we find that estimating σ_e causes the local optima problem. Therefore, we use a grid search method by setting it at different levels and choose the one that gives the maximum likelihood value.[13]

4.4.2. Postclick Utility δ_{ij}. The identification of the postclick seller fixed effects ω_j comes from the heterogeneity in conversion rates across sellers. Some sellers do not have a very high click rate but a very high conversion rate conditional on clicking (Figure 1). This is interpreted as the seller having a high value ω_j on the seller page. Intuitively, ω_j is identified from the deviation of the actual conversion rate from the expected if the realized value after search is i.i.d. across consumers and products (when $\delta_{ij} = \epsilon_{ij}$). This is different from a choice model setting, where one final choice from consumers is observed. In a search model, we observe two-stage choices from consumers—the

clicking decision and purchase decision conditional on clicking. In the model, we have imposed exclusion restrictions on the information structure during the two-stage choices: the postclick seller fixed effects ω's (parameters in the postclick utility δ_{ij}) do not influence the consumer clicking choice but only affect the purchase decisions after search;[14] the preclick utility V_{ij}, on the other hand, is observed prior to clicking and does not affect δ_{ij}, which is only realized after search. Observing the two-stage choices is essential in separating δ_{ij} from V_{ij}. Recall from Section 3.1 that the average ω_j across sellers is normalized to 0. The variance of the idiosyncratic term ϵ_{ij} is normalized to 1 (Kim et al. 2010).

4.4.3. Search Cost c_i. How consumers search depends on both the purchase utility and the search cost. Conditional on the utility parameters, the average number of searches in data identifies the average search cost $\bar{c}$ in Equation (4). The heterogeneity in the number of searches among consumers identifies the variance σ_c^2. We impose the exclusion restriction as in Chen and Yao (2016) by assuming that the observed characteristics in X_{ij} will only affect V_{ij} (and thus the actual utility u_{ij}) but not the search cost. This is supported by Table 3, where we show that X's are correlated with the conversion rate but not the number of searches.

4.5. Monte Carlo Study
We use a Monte Carlo simulation to demonstrate that model parameters can be recovered using our estimation method. We generate a data set with 100,000 consumers searching among five sellers. The consumers' search and purchase behaviors are simulated based on assumed model parameters. We further draw three covariates: price, star level, and same location for each seller. It is important to note that the search patterns are close between the simulated data and true data. For the simulated data, 89.4% consumers search once, 8.6% search twice, 1.7% search three times, and the rest search four or more times. It resembles the observed search history as reported in Table 1.

During the estimation, we draw random variables c_i, e's, ϵ's from their respective distributions. We make 500 sets of draws using the GHK algorithm for each consumer as described in Section 4.3. The estimation results for the Monte Carlo study are reported in Table 4. Column (1) shows the true parameters we use in the simulation.[15] The starting values for the estimation are set at 0 for all parameters (column (2)). Column (3) shows the estimated values and their standard errors. The estimated values are very close to the true values despite a set of faraway starting values. This Monte Carlo exercise suggests that the proposed model can be identified from data with the

Table 4. Monte Carlo Simulation Results

Parameters	True values	Starting values	Estimated parameters	Restrict $\omega's = 0$
	(1)	(2)	(3)	(4)
Utility parameters				
Constant	−2.9	0	−2.9158 (0.0046)	−2.9060 (0.0041)
High price	−0.1	0	−0.1017 (0.0023)	−1.0001 (0.0023)
High star	0.1	0	0.1031 (0.0023)	0.1032 (0.0023)
Same location	0.2	0	0.2074 (0.0030)	0.1995 (0.0031)
Search cost parameters				
Mean of log search cost	−5.0	0	−5.0356 (0.0071)	−5.1161 (0.0075)
Std. dev. of log search cost	0.5	0	0.4853 (0.0081)	0.4393 (0.0087)
Unobserved attributes				
ξ_1	−0.1	0	−0.0996 (0.0014)	−0.0965 (0.0014)
ξ_2	0.1	0	0.0996 (0.0014)	0.0965 (0.0014)
ξ_3	−0.1	0	−0.1041 (0.0018)	−0.0977 (0.0018)
ξ_4	0.1	0	0.0996 (0.0014)	0.0965 (0.0014)
ξ_5	0.0	0	0.0045 (0.0023)	0.0012 (0.0023)
ω_1	0.1	0	0.1042 (0.0121)	—
ω_2	−0.15	0	−0.1483 (0.0130)	—
ω_3	0.15	0	0.1576 (0.0115)	—
ω_4	−0.1	0	−0.1076 (0.0154)	—
ω_5	0.0	0	−0.0059 (0.0262)	—
LL			−225,319.5	−225,397.2
AIC			450,663.0	450,810.4
BIC			450,777.2	450,886.5

Notes. Standard errors are reported in parentheses. LL, log likelihood.

observed level of search activities and that the GHK method is able to recover the true model parameters in the estimation.

After showing that the proposed model can be identified from the search and purchase data, we further zoom in on the effect of including the postclick seller fixed effects $\omega's$ in the model, which represent the value of information on the seller page that is shared across consumers. To show the impact of the $\omega's$, we estimate a version of the model that assumes $\omega's = 0$ for all sellers, whereas in the actual data generating process, the sellers have their unique value of information on the seller page, as specified in the Monte Carlo study (Table 4, column (1)).

The estimation results from the restrictive model are reported in column (4) The model fit statistics, including both the Akaike information criterion (AIC) and Bayesian information criterion (BIC), are worse in the restrictive model. Although the parameter estimates are reasonably close to the true values, they are not as good as the true model with the $\omega's$, especially for the search cost parameters. We simulate the consumer search and purchase behavior separately using the estimated parameters in the full model and the restrictive model. The two models imply very different conversion rate patterns. The proposed model predicts that the conversion rates are higher for sellers with larger $\omega's$, such as sellers 1 and 3, and lower for those with lower $\omega's$, such as sellers 2 and 4 (Figure 4). However, if we restrict $\omega's = 0$, the predicted conversion

rates do not show such patterns. The Monte Carlo exercise shows that the postclick seller fixed effects $\omega's$ can be identified from data. They capture the value of information on the seller page, which is reflected by the conversion rates after clicking across sellers.

5. Empirical Application and Results

We estimate the proposed model using data from Taobao. On the search results page, consumers can observe the price, seller star level, and location of the seller. Therefore, we include these attributes in the portion of the utility function that is observed prior to search. We also observe some consumer

Figure 4. (Color online) Conversion Rate Patterns from Different $\omega's$
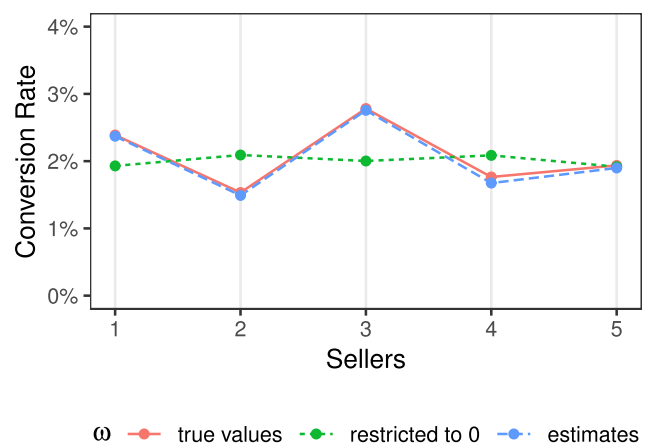
characteristics—age, gender, and the number of past transactions (as a proxy for her affinity of buying from Taobao)—as well as whether it is during weekend and working hours when the consumer conducts the search. Table 3 shows that across consumers with different observed characteristics, the number of searches is similar but the conversion rate varies. These observed consumer characteristics may affect the baseline utility toward purchasing an iPad from Taobao. Therefore we include them in the utility function.

Table 5 reports the estimation results. The coefficient for the high-price group is negative and significant. The average price difference between the two price groups is ¥232. Therefore, the price coefficient is −0.040 for ¥100. We compare the estimate with the previous literature on consumer search. Using data from Expedia, Ursu (2018) estimates the price coefficient to

be in the range of −0.28 and −0.12 for US$100. Converting these coefficients using Chinese yuan, the range is −0.045 and −0.019 for ¥100. The lower bound is very close to our estimate. For robustness, we also estimate another model by categorizing the sellers into three price groups—low-, medium-, and high-price—and obtain a similar implied price coefficient (see Online Appendix 2 for details). The coefficient for the high-star-level group is positive and significant. Similarly, the model in the robustness check categorizes the sellers into three star groups. The results also suggest that consumers are more likely to search and buy from sellers living in the same geographical location. Furthermore, older and male consumers tend to have a higher purchase utility. The coefficient for those missing demographics is positive and significant, but because age is coded as 0 for these consumers,

Table 5. Estimation Results

Parameters	Estimates (std. error)	
Utility parameters		
Constant	−3.5247 (0.0062)	
Price	−0.0932 (0.0027)	
High star	0.2234 (0.0044)	
Same location	0.2211 (0.0043)	
Age (log)	0.0222 (0.0020)	
Male	0.0159 (0.0045)	
Missing demographics	0.0668 (0.0042)	
Number of past transactions (log)	0.0058 (0.0007)	
Working hours	0.0255 (0.0008)	
Weekend	−0.0022 (0.0018)	
Search cost parameters		
Mean of log search cost	−4.7045 (0.0023)	
Std. dev. of log search cost	0.5632 (0.0012)	
Estimated ξ's and ω's	Preclick ξ_j	Postclick ω_j
Seller 1	0.7790 (0.0031)	−0.5288 (0.0111)
Seller 2	0.5185 (0.0036)	−0.2209 (0.0137)
Seller 3	0.2045 (0.0038)	0.0501 (0.0087)
Seller 4	0.0712 (0.0042)	0.2563 (0.0096)
Seller 5	0.3751 (0.0047)	−0.3320 (0.0151)
Seller 6	−0.0233 (0.0054)	−0.6489 (0.0141)
Seller 7	0.0410 (0.0053)	−0.0646 (0.0079)
Seller 8	0.1271 (0.0055)	0.1808 (0.0073)
Seller 9	−0.1979 (0.0059)	−0.0215 (0.0075)
Seller 10	−0.2163 (0.0062)	0.1857 (0.0074)
Seller 11	−0.1901 (0.0062)	−0.2701 (0.0074)
Seller 12	−0.0877 (0.0069)	0.2439 (0.0076)
Seller 13	0.0074 (0.0069)	0.2674 (0.0078)
Seller 14	−0.2497 (0.0065)	0.4807 (0.0080)
Seller 15	−0.1262 (0.0074)	0.0495 (0.0080)
Seller 16	−0.1234 (0.0078)	0.1283 (0.0081)
Seller 17	−0.3442 (0.0173)	0.1330 (0.0080)
Seller 18	−0.1056 (0.0120)	0.1919 (0.0074)
Seller 19	−0.1077 (0.0207)	0.0693 (0.0073)
Seller 20	−0.3518 (0.0172)	−0.1500 (0.0416)
Observations	104,819	
LL	−287,988.3	
AIC	576,072.6	
BIC	576,115.3	

Note. LL, log likelihood.

they actually tend to have a lower utility on average. Consumers with a larger number of past transactions on Taobao have a higher utility toward buying iPad from Taobao. Consumers who search during working hours are likely to have a higher utility, but search during weekends does not make a significant difference.

We assume the search costs across consumers follow a log-normal distribution (Equation (4)). Taking the exponential of the estimates, the average search cost is 0.0106 and the standard deviation is 0.0065. Turning the search cost into a dollar amount, we divide the average search cost by the price coefficient and find the average search cost to be ¥26.4 (about $4.3).

Table 5, column (1) reports the estimated preclick fixed effects ξ_j (Equation (2)) for each seller. Results show a large difference across sellers, which contributes to the different clicking rates across sellers. Column (2) reports the estimated postclick fixed effects ω_j (Equation (3)), which represent the value of information on the seller page. Sellers with a higher ω_j are more likely to convert browsers into purchasers. Comparing the two columns, sellers with a high ability of attracting clicks do not necessarily have a high ability of conversion. For example, seller 1 has the highest ξ but a very low ω, explaining why his conversion rate is lower than many others. By contrast, seller 14 has a low ξ but the highest ω. Indeed, the correlation between $\xi's$ and $\omega's$ is close to 0 across sellers.

The estimated preclick seller fixed effects $\xi's$ include the impact from ranking and the seller quality, and the value can change when the ranking changes. The important thing is that the estimated $\xi's$ capture the inferred value of the preclick information during the time period when the consumers shop, which is informative for retargeting consumers who searched during that time period.

In the data, only 11.1% of consumers searched for more than one seller, and 2.2% searched for more than two. The lack of search can hurt sales, as sellers with high ξ (e.g., seller 1) draw most of clicks but cannot convert, whereas sellers with high ω but low ξ (e.g., seller 14) cannot attract consumers to click. To illustrate the impact, we use the estimation results to simulate consumer searches and purchases, assuming that $\omega's$ are fully revealed on the search results page (although consumers still have to search ϵ_{ij} in the utility function). Compared with the current scenario, the conversion rate will increase by 11%. The large increase demonstrates the value of revealing the sellers' information to consumers.

As a test of the face validity of the results, we estimate the model using data from another product, the iPad 4 16G WiFi, that was sold on Taobao during the same sample period. The estimation results are reported in Online Appendix 3. The results are similar to those in Table 5. This is what we expect if the model correctly specifies consumers' behaviors, as individuals on the same retail platform searching for similar products should not be too different. We view the result as indirect support for the proposed model.

6. The Effectiveness of Retargeting

The model estimation recovers structural parameters including the consumer utility function and the search cost. With the estimated model, we are able to simulate consumer search and purchase behaviors and infer the preference of consumers. We use the results to investigate how different consumers would behave under various counterfactual retargeting policies. Retargeting can potentially have a substantial profit impact because, in our data, 98.7% of consumers who search for the product end up not buying it. These consumers represent a great asset for sellers and Taobao because they have expressed a purchase interest. To implement the proposed retargeting policies, Taobao can rely on direct messaging on its platform (through its mobile app or website) to reach customers. Just as how Taobao can send coupons or communication messages, it is in a good position to send retargeting messages on behalf of participating sellers.

From the estimation results, we simulate the search and purchase decisions of all consumers in the data. Those who do not purchase will be retargeted. We use 1,000 simulated draws of $e's$ in Equation (2) and in Equation (3) for each consumer and seller pair. To calculate the profit, we assume that Taobao takes 15% of the seller revenue as the referral fee[16] and the profit margin for all sellers is 10% after the referral fee. This is equal to an average profit of ¥220 for sellers and ¥330 for Taobao. Furthermore, we assume that the expected utility and the realized utility remain unchanged when a consumer is retargeted.[17]

6.1. Retargeting Consumers with Coupons

This counterfactual investigates the effectiveness of retargeting consumers by offering coupons that can be redeemed upon purchase. We assume that the value of a coupon is ¥110, about 5% of the average selling price or half of the sellers' profit.[18] The value of the coupon is the same for every consumer, and the cost of the coupon is borne by sellers. We also assume that, if consumers respond to the offer, they have to pay the search cost to learn other details from the seller.

As a benchmark, we first simulate the responses when, for every consumer, a seller is randomly chosen to offer the coupon. The first row in Table 6 reports the average conversion rate and the average value of a coupon, which is the expected profit for sellers after taking account of coupon cost. The conversion rate is only 0.008%. Compared with the 1.35% simulated average purchase rate before retargeting, the effectiveness

Table 6. Results from Random Retargeting with Coupons and the Heterogeneity

	Conversion rate of retargeting (%)	Expected value (¥)
Random retargeting for all consumers w/o purchase	0.0083	0.0089
Heterogeneity based on observed consumer characteristics		
Past number of transactions		
0–30	0.0074	0.0080
31–150	0.0080	0.0087
151 or above	0.0087	0.0090
Gender		
Male	0.0090	0.0097
Female	0.0084	0.0091
Age		
18–25	0.0084	0.0090
26–30	0.0089	0.0097
31 or above	0.0092	0.0099
Time of visit		
Working hours	0.0087	0.0093
Nonworking hours	0.0077	0.0083
Day of visit		
Weekday	0.0083	0.0090
Weekend	0.0082	0.0088
Heterogeneity based on consumer search behaviors		
Own or competitors' customers		
Send to own customers	0.1040	0.1113
Send to competitors' customers	0.0025	0.0027
Search intensity		
Searched one seller	0.0059	0.0063
Searched two sellers	0.0257	0.0278
Searched three or more sellers	0.0347	0.0377
Order of search (among own customers)		
Searched two sellers		
First searched seller	0.2342	0.2514
Second searched seller	0.1677	0.1828
Searched three sellers		
First searched seller	0.2480	0.2556
Second searched seller	0.1688	0.1834
Last searched seller	0.1422	0.1543
Heterogeneity based on seller characteristics		
Sales volume		
High	0.0146	0.0157
Low	0.0019	0.0021
Click rate		
High	0.0143	0.0152
Low	0.0023	0.0027
Conversion rate		
High	0.0102	0.0112
Low	0.0063	0.0066

of such a random retargeting policy is rather low, and consequently, the average value for every coupon is merely ¥0.009.

We further investigate the heterogeneity in consumer responses. The second panel of Table 6 shows that the conversion rates for male and older consumers who made a large number of purchases at Taobao before are higher than those of the others. In addition, those who visited Taobao during working hours are also more responsive to retargeting. The heterogeneity between consumers with different characteristics, however, is just moderate.

Substantial heterogeneity arises when we compare consumers with different types of search behaviors. The results are reported in the third panel of Table 6. For a seller, the conversion rate from his or her "own" customers (i.e., consumers who have searched the seller's page before) is 0.104%, more than 40 times higher than the conversion rate from competitors' customers (i.e., consumers who have only clicked on other sellers' links). The reason for the difference is that if a consumer chose to click on the seller's link before, she would have a high expected utility from the seller. Although (before retargeting) the realized

utility of buying is lower than the no-purchase option, the likelihood that the consumer is attracted by the seller's coupon offer is still much larger than when the coupon is offered by sellers with whom she has low expected utility. Because of the difference, the value for each coupon sent to the seller's own customers is ¥0.11, and the value of sending to other consumers is negligible.

Responses are also heterogeneous among consumers who exhibited different search intensity in terms of the number of sellers they have searched. Results show that consumers with high search intensity have a higher conversion rate from retargeting. Although consumers with a lower search cost can choose to search more, on average, consumers with intensive search activities are more likely to have a high valuation for the product.

Furthermore, the order of search is another useful metric to predict the response. Among consumers who have clicked on multiple links, the conversion rate is the highest if they clicked on the seller's link first and monotonically decreasing in the search order. For example, for a consumer who has searched three sellers, the conversion rate if the coupon is sent by the seller searched first is 74% higher than if it is sent from the seller who was searched last. The reason for the decreasing conversion rate is that the consumer ordered which seller to look at based on the expected utility. Note that this is different from the U-shaped conversion rate exhibited in Figure 3. As consumers will make a purchase (and therefore stop searching) when they find a good offering, the conversion rate for the last clicked option in the search process is typically high. By contrast, if they abandon the search without making a purchase, their expected utility for the last option is lower than for the first option. This subtle difference indicates that we should not predict the response to retargeting based only on those who have made purchases and ignore the systematic difference between the two groups of consumers.

The effectiveness of retargeting also varies across sellers with different characteristics. We segment sellers into high and low groups based on their sales volume, click rates, and conversion rates before retargeting. The last panel of Table 6 shows that retargeting is more effective for the groups with high sales, click rates, and conversion rates than their counterparts. For example, the expected conversion rate from retargeting is more than nine times higher among sellers with a high number of clicks compared with sellers with low click numbers.

The aforementioned results suggest the large heterogeneity across consumers and sellers in terms of the retargeting effectiveness. Some consumers have "higher value" in retargeting such as consumers with higher search intensity. The retargeting effectiveness

also depends on the particular seller-consumer pair. For example, a consumer who has searched from a particular seller is more valuable for that seller than others when retargeting with a coupon. Therefore, a better matching between consumers and sellers is needed to improve the efficiency of retargeting.

We now investigate how the online retail platform can run effective retargeting by considering several counterfactual policies. The first uses random retargeting as a benchmark. The second considers a typical industry practice that retargets consumers with the option that has been searched most recently (Cristal 2015). Under the third policy, Taobao uses a pricing mechanism (second-price auctions) to make sure sellers with a higher chance to convert consumers will self-select to participate in retargeting. In the fourth policy, in addition to the second-price auction mechanism, Taobao also provides sellers with the information on the search behaviors of each consumer, which can help sellers achieve a better match with consumers in retargeting.

Results for these counterfactual policies are reported in Table 7. Column (1) shows the conversion rate, total consumer welfare, and the total profit for the platform and sellers, taking account of the coupon cost, under the random matching policy as a benchmark. As the conversion rate is only 0.008%, the consumer welfare and profit generated from retargeting coupons are quite small.

Column (2) reports results from another benchmark case in which the coupon is sent from the seller whose link was clicked last during the search process. It shows that the conversion rate will increase 10-fold from random retargeting. As a result, the consumer welfare and the total profit will also increase significantly. These increases are because consumers would only click on the seller's link if they have a high expected utility. Coupons sent from the seller therefore will have a conversion rate much higher than that from a random seller.

If every seller sends messages to all consumers who have visited his site, consumers will be annoyed because they will receive way too many messages from different sellers. The messages can also "crowd out" one another, as such consumers may simply ignore them. We therefore consider a counterfactual scenario under which sellers compete for the right to send coupons through second-price auctions. We assume that each seller has full knowledge on the model parameters, but he only knows the distribution of e's and ϵ's of each individual consumer. With a second-price auction, at the equilibrium, the optimal strategy for the seller is to bid according to the expected profit (after integrating out the e's and ϵ's) from sending the coupon. A seller whose offering has the highest expected utility will bid the most and

Table 7. Counterfactual—Retargeting with Coupons

	Random retargeting	Retargeting by the last click	Retargeting through second-priced bids: No consumer information	Retargeting through second-priced bids: With consumer search history
Conversion rate (%)	0.008	0.086	0.041	0.098
Consumer welfare (¥)	637	5,108	2,459	5,797
Seller profit (¥)	—	—	667	9,645
Platform profit (¥)	—	—	17,151	34,337
Total profit (¥)	3,760	38,387	17,818	43,982

pay the expected profit of the second-highest seller. With the auction mechanism, consumers are more likely to receive coupons from sellers with a higher probability to convert them. It can help narrow down which seller the platform should help sending the retargeting coupon and avoid potential negative responses from the retargeting effort. Therefore, the auction mechanism is expected to improve the efficiency of retargeting.

Column (3) in Table 7 reports the results. Note that in this counterfactual scenario, sellers have no other information about this consumer other than that the target consumer has searched on Taobao without making a purchase. In particular, they do not have information on how the consumer searched. The conversion rate is higher than that in column (1), when the retargeting comes from random sellers, suggesting that the auction mechanism increases the efficiency of retargeting. However, the conversion rate is significantly lower than that in column (2), when coupons are offered by sellers whose links were clicked last. The consumer welfare and the total profit for Taobao and sellers are also lower than column (2). These results suggest that the benefits from being given one single piece of information on how consumers searched (i.e., which link was clicked the last) dominate the benefits from sellers' self-selection through the auction mechanism.

Table 6 shows that the conversion rate is monotonically decreasing in the search order. This suggests that there may be a better match for consumers based on search history that can further improve the efficiency of retargeting. Given the value of consumers' search behaviors, we study the last scenario in which Taobao provides sellers with the full search history of consumers, including the search intensity and the order of search. Sellers can better predict how consumers with different search behaviors will respond and can adjust their bids to retargeting the consumers accordingly. Column (4) shows that the conversion rate increases to 0.1%, significantly higher than all other scenarios. In addition, the consumer welfare, seller's profit, and Taobao's profit all increase substantially, suggesting that providing the information benefits all parties in retargeting. The increase in retargeting effectiveness is because the information revealed during search enables a better match between consumers and sellers for retargeting. For example, consumers are more likely to receive a retargeting coupon from sellers who they click on first, as well as from sellers with higher conversion rates. In this scenario, consumers are more likely to respond to the retargeting coupon compared with the previous cases. To summarize, providing sellers with the information on consumers' search history will significantly improve the effectiveness of retargeting.

6.2. Retargeting Consumers with Seller Recommendations

The ¥44K total profit in column (4) of Table 7 represents an approximately 6% increase from the original profit without retargeting. Such a gain is economically meaningful but not too substantial. We investigate in this subsection whether there is a strategy that can further enhance the effectiveness of retargeting.

We have discussed in Section 5 that purchases will significantly increase if consumers have full information on the exact seller offerings. The counterfactual exercise in this subsection thus considers providing retargeted consumers with the sellers' information. One way to implement this idea is to recommend a seller through personalized messages (e.g., direct messaging on the Taobao platform). We assume that the recommendation reveals the actual seller offering. For example, the message can include a detailed description of product features, consumer reviews, and warranty, for example. After receiving the recommendation, the consumer will buy from the seller if the actual purchase utility is larger than the no-purchase value. We assume that if the consumer responds to the recommendation, she still has to pay the search cost to complete the transaction process on the seller's page.

We first investigate the scenario in which Taobao randomly recommends one seller to each consumer. Column (1) in Table 8 shows that the conversion rate is 0.1%, which is close to the most effective practice in the previous counterfactual exercise when sellers offer coupons (see column (4) in Table 7). The increase in the consumer welfare and the total profit is also significant. Results suggest that providing consumers with the sellers' information is an effective retargeting strategy, even when sellers are chosen randomly.

Table 8. Counterfactual—Retargeting by Seller Recommendation

	Random retargeting	Retargeting through second-priced bids: No consumer information	Retargeting through second-priced bids: With consumer search history
Conversion rate (%)	0.120	0.342	0.359
Consumer welfare (¥)	33,095	96,674	101,392
Seller profit (¥)	—	2,347	8,204
Platform profit (¥)	—	181,529	186,726
Total profit (¥)	68,783	183,876	194,930

Similar to the previous exercise, we show that the retargeting effectiveness increases when sellers bid to be recommended for each consumer based on second-price auctions. The auction mechanism can help achieve a better matching between sellers and consumers. Column (2) of the table shows that the conversion rate will increase to 0.34%, almost three times as much as under random retargeting. The consumer welfare and the total profit for Taobao and sellers are also higher. Again, this result demonstrates that self-selection among sellers via the auction mechanism helps increase the efficiency of retargeting.

Finally, column (3) of the table provides results when Taobao further provides sellers with the information on consumers' search behaviors. The search history helps sellers choose consumers who have a higher chance of responding to the recommendation. We find that the conversion rate, consumer welfare, and the total profit further increase from that in column (2), showing the significant benefits of allowing sellers to target consumers based on their search history.

Comparing Tables 7 and 8, we show that seller recommendation is more effective than offering coupons as a retargeting strategy. We also find that the utility gains from consumers who respond to seller recommendations are higher than those from consumers who redeem coupons. What is the reason behind these findings? We find the top three sellers who win the most bids for sending coupons are sellers 1, 2, and 3. These sellers have high ξ's and therefore high click rates. By contrast, the top three sellers who win the most bids to become the recommended seller are sellers 4, 3, and 14. These sellers have the highest actual utility averaged across consumers but, compared with sellers 1 and 2, have a lower ability of attracting consumers to click on their links. The seller recommendation helps reveal their actual offerings, and therefore, consumers can find a better match rather than buying from sellers they have searched before. This comparison illustrates the importance of providing consumers with the information of recommended sellers. Note that the ¥195K total profit in column (3) represents a 25% increase from the original profit without retargeting. This suggests that, if done

in the right way, retargeting can be a very effective marketing strategy.

To summarize, we have shown in the counterfactuals that the consumers' search behaviors are valuable in predicting the effectiveness in retargeting. This is because we can infer the preference of the consumers based on their search history, more so than other observed consumer characteristics, such as demographic information. Thus providing sellers the information on consumer search history greatly improves the retargeting effectiveness. Online retail platforms can utilize the consumer search data to increase own profit. The search information also benefits consumers and participating sellers.[19]

7. Conclusions and Limitations

This paper studies the value of consumer search history in retargeting consumers who have searched online but did not purchase. To infer the consumer preference from search history, we build and estimate a structural sequential search model. For estimation, we propose a GHK-type estimator by making recursive draws from truncated distributions that arise from the observed search and choice behaviors. The proposed estimator directly calculates the likelihood function and achieves a smooth likelihood without the need to specify scaling factors.

Because the online retail platform did not run retargeting, we run a series of counterfactuals to explore the effectiveness of retargeting based on different strategies. There are several key findings. First, we show a substantial heterogeneity in responses to retargeting among consumers who exhibited different search behaviors. Second, recommending a seller can be more effective in increasing the conversion rate than offering coupons. Third, platforms can use a pricing mechanism such as auctions to improve the retargeting efficiency. Finally, the retargeting effectiveness further improves when sellers have access to the consumer search history, and both sellers and consumers will benefit.

Our study has several limitations. The measured effects are based on several model assumptions. For example, we assume that the consumer need for iPad

still exists when reached by retargeting actions, and consumers do not update their expectations on sellers from the retargeting actions and are nonstrategic in the response. Because retargeting has not been adopted in our empirical context, we are unable to study whether the above-mentioned assumptions are valid and how they may influence the results. We also abstract away from any potential backlash from sellers if Taobao monetizes the retargeting service or any consumer annoyance or privacy concerns when retargeted. We view the counterfactual exercise as a thought experiment that can provide useful guidance for firms in the early stage of planning retargeting strategies. The findings can be further tested with actual field experiments using a broader range of product categories and in different countries for the generalizability. Finally, although we show the significance of considering consumer search history, the study has abstracted away from other important aspects for retargeting, including the timing of serving ads or how often to serve ads (e.g., Sahni et al. 2019) and how consumers respond to competing ads and spillover effects (e.g., Sahni 2016). Future research should further investigate how these other aspects will impact the efficiency of retargeting.

Acknowledgments

Appendix
A.1. Calculate Reservation Utility with Newton's Method

Recall that solving for the normalized reservation utility is equivalent to finding the solution for the function $g(\zeta_i) = 0$, where $g(\zeta) = (1 - \Phi(\zeta))(\frac{\phi(\zeta)}{1-\Phi(\zeta)} - \zeta) - x$. To calculate the reservation utility using Newton's method, we iterate ζ according to the following equation:

$$\zeta_{k+1} = \zeta_k - \frac{g(\zeta_k)}{g'(\zeta_k)} \quad (A.1)$$

until ζ_{n+1} and ζ_n are very close (see Section 4.1). Subscript i is omitted for simplicity. In this appendix, we show the derivation of the next guess ζ_{k+1} in the iteration process.

We start by calculating the first-order derivative of the function $g(\zeta_n)$:

$$g'(\zeta_k) = -\phi(\zeta_k) \cdot \left(\frac{\phi(\zeta_k)}{1 - \Phi(\zeta_k)} - \zeta_k \right) + (1 - \Phi(\zeta_k))$$

$$\cdot \left(\frac{\phi'(\zeta_k)}{(1 - \Phi(\zeta_k))} - \frac{\phi(\zeta_k) \cdot (-\phi(\zeta_k))}{(1 - \Phi(\zeta_k))^2} - 1 \right)$$

$$= -\frac{\phi^2(\zeta_k)}{1 - \Phi(\zeta_k)} + \phi(\zeta_k) \cdot \zeta_k + \phi'(\zeta_k) + \frac{\phi^2(\zeta_k)}{1 - \Phi(\zeta_k)} - (1 - \Phi(\zeta_k))$$

$$= \phi(\zeta_k) \cdot \zeta_k + \phi(\zeta_k) \cdot (-\zeta_k) - (1 - \Phi(\zeta_k))$$

$$= \Phi(\zeta_k) - 1.$$

We plug $g'(\zeta_k) = \Phi(\zeta_k) - 1$ into Equation (A.1):

$$\zeta_{k+1} = \zeta_k - \frac{g(\zeta_k)}{g'(\zeta_k)}$$

$$= \zeta_k - \frac{(1 - \Phi(\zeta_k)) \cdot \left(\frac{\phi(\zeta_k)}{1-\Phi(\zeta_k)} - \zeta_k \right) - x}{\Phi(\zeta_k) - 1}$$

$$= \zeta_k + \frac{\phi(\zeta_k)}{1 - \Phi(\zeta_k)} - \zeta_k + \frac{x}{\Phi(\zeta_k) - 1}$$

$$= \frac{\phi(\zeta_k) - x}{1 - \Phi(\zeta_k)}.$$

Therefore, the next guess for ζ_{k+1} has a simple form:

$$\zeta_{k+1} = \frac{\phi(\zeta_k) - x}{1 - \Phi(\zeta_k)}.$$

A.2. Simulated Maximum Likelihood Using the GHK Method

In this appendix, we describe the sampling procedure of the GHK simulator for purchasers (see Section 4.3 for non-purchasers). We further separate out consumers who purchase the last clicked option and those who purchase an option that is not the last clicked. This is because the likelihood function can be simplified differently for these two types of purchase behaviors.

A.2.1. Purchase the Last Clicked Option ($j^* = J$). For consumers who purchase the last clicked option, the search conditions from Equations (7)–(10) can be simplified as follows:
- The order condition (Equation (7)):

$$\bar{V}_1 + e_1 > \bar{V}_2 + e_2 > \cdots > \bar{V}_J + e_J$$
$$> \max_{j' \notin (1,\dots,J)} (\bar{V}_{j'} + e_{j'}).$$

- The continue to search condition when $J > 1$ (Equation (8)):

$$\max \left(0, \max_{j \in (1,\dots,J-1)} (\bar{V}_j + e_j + \omega_j + \epsilon_j) \right)$$
$$< \bar{V}_J + e_J + \sigma \cdot \zeta(x).$$

- The stop searching condition (Equation (9)):

$$\bar{V}_J + e_J + \omega_J + \epsilon_J > \max_{j' \notin (1,\dots,J)} (\bar{V}_{j'} + e_{j'}) + \sigma \cdot \zeta(x).$$

- The choice condition:

$$\bar{V}_J + e_J + \omega_J + \epsilon_J > \max \left(0, \max_{j \in (1,\dots,J-1)} (\bar{V}_j + e_j + \omega_j + \epsilon_j) \right).$$

With the above-mentioned search conditions, we describe the detailed sampling procedure:

1. Draw $c^{(s)}$ from the log-normal distribution (Equation (4)) and calculate the normalized search cost $x^{(s)} = \frac{c^{(s)}}{\sigma}$.

2. Draw $e_{j'}$'s for the unsearched options with no bounds.

3. Sample e_J for the last searched option, whose truncated region comes from the order condition and part of the continue to search condition (if $J > 1$).[20] The truncated region depends on the $e_j^{(s)}$'s draws.

   a. Draw $e_J^{(s)}$ with lower bound $\underline{b} = \max(-\bar{V}_J - \sigma \cdot \zeta(x),$ $\max_{j' \notin (1,\ldots,J)}(-\bar{V}_J + \bar{V}_{j'} + e_{j'}^{(s)}))$.

4. Sample $e_{J-1}, \ldots, e_1$ recursively (if $J > 1$). The draws need to satisfy the order condition, which depends on the previous draws of $e_j^{(s)}$.

   a. Recursively draw $e_j^{(s)}$ from its distribution with lower bound $\underline{b} = \bar{V}_{j+1} + e_{j+1}^{(s)} - \bar{V}_j$, $\forall\, j = J-1, \ldots, 1$. We get $z_j^{(s)} = \bar{V}_j + e_j^{(s)} + \sigma \cdot \zeta(x^{(s)}), \forall j = 1, \ldots, J$.

5. Sample ϵ_j, $j = 1, \ldots, J-1$ (if $J > 1$). The draws need to satisfy the continue to search condition, which depends on the previous draws of $e_j^{(s)}$, $j = 1, \ldots, J$.

   a. Draw $\epsilon_j^{(s)}$, $j = 1, \ldots, J-1$ with upper bound $\bar{b} = -\bar{V}_j - e_j^{(s)} - \omega_j + \bar{V}_J + e_J^{(s)} + \sigma \cdot \zeta(x)$. We get $u_j^{(s)} = \bar{V}_j + e_j^{(s)} + \omega_j + \epsilon_j^{(s)}$, $j = 1, \ldots, J-1$.

6. We then evaluate the likelihood of the stop searching condition and the choice condition. The probability that ϵ_J satisfy the inequality can be directly evaluated.

   a. The lower bound is $\underline{b} = -\bar{V}_J - e_J^{(s)} - \omega_J + \max(0, \max_{j' \notin (1,\ldots,J)} z_{ij'}, \max_{j \in (1,\ldots,J-1)} u_j^{(s)})$. The probability that ϵ_J satisfy the inequality is $P(\epsilon_J > \underline{b}) = 1 - \Phi(\underline{b})$.

7. Finally, the likelihood can be approximated by the average of the probabilities using a set of draws:

$$L^s = \frac{1}{S} \sum_{s=1}^{S} \left\{ \left(1 - \Phi^e\left(-\bar{V}_J - \sigma \cdot \zeta(x), \max_{j' \notin (1,\ldots,J)}\left(-\bar{V}_J + \bar{V}_{j'} + e_{j'}^{(s)}\right)\right)\right) \right.$$

$$\cdot \prod_{j=1}^{J-1} \left(1 - \Phi^e\left(\bar{V}_{j+1} + e_{j+1}^{(s)} - \bar{V}_j\right)\right) \cdot$$

$$\prod_{j=1}^{J-1} \Phi\left(-\bar{V}_j - e_j^{(s)} - \omega_j + z_J^{(s)}\right) \cdot$$

$$\left. \left(1 - \Phi\left(-\bar{V}_J - e_J^{(s)} - \omega_J + \max\left(0, \max_{j' \notin (1,\ldots,J)} z_{ij'}, \max_{j \in (1,\ldots,J-1)} u_j^{(s)}\right)\right)\right)\right\}.$$

A.2.2. Purchase an Option That Is Not the Last Clicked ($j^\star \neq J$).

For consumers who purchase an option that is not the last clicked, the search conditions from Equations (7)–(10) can be simplified as follows:

- The order condition (Equation (7)):

$$\bar{V}_1 + e_1 > \bar{V}_2 + e_2 > \ldots > \bar{V}_J + e_J$$
$$> \max_{j' \notin (1,\ldots,J)}\left(\bar{V}_{j'} + e_{j'}\right).$$

- The continue to search condition (Equation (8)):

$$\max\left(0, \bar{V}_{j^\star} + e_{j^\star} + \omega_{j^\star} + \epsilon_{j^\star}\right) < \bar{V}_J + e_J + \sigma \cdot \zeta(x).$$

- The stop searching condition (Equation (9)):

$$\bar{V}_{j^\star} + e_{j^\star} + \omega_{j^\star} + \epsilon_{j^\star} > \max_{j' \notin (1,\ldots,J)}\left(\bar{V}_{j'} + e_{j'}\right) + \sigma \cdot \zeta(x).$$

- The choice condition:

$$\bar{V}_{j^\star} + e_{j^\star} + \omega_{j^\star} + \epsilon_{j^\star}$$
$$> \max\left(0, \max_{j \in (1,\ldots,J) \setminus j^\star}\left(\bar{V}_j + e_j + \omega_j + \epsilon_j\right)\right).$$

With the above-mentioned search conditions, we describe the detailed sampling procedure.

1. Draw $c^{(s)}$ from the log-normal distribution (Equation (4)) and calculate the normalized search cost $x^{(s)} = \frac{c^{(s)}}{\sigma}$.

2. Draw $e_{j'}$'s for the unsearched options with no bounds.

3. Sample e_J for the last searched option, whose truncated region comes from the order condition. The truncated region depends on the $e_j^{(s)}$'s draws.

   a. Draw $e_J^{(s)}$ with lower bound $\underline{b} = \max(\max_{j' \notin (1,\ldots,J)}$ $(-\bar{V}_J + \bar{V}_{j'} + e_{j'}^{(s)}), -\bar{V}_J - \sigma \cdot \zeta(x^{(s)}))$.

4. Sample $e_{J-1}, \ldots, e_1$ recursively. The draws need to satisfy the order condition (Equation (12)), which depends on the previous draws of $e_j^{(s)}$.

   a. Recursively draw $e_j^{(s)}$ from its distribution with lower bound $\bar{V}_{j+1} + e_{j+1}^{(s)} - \bar{V}_j$, $\forall\, j = J-1, \ldots, 1$. We get $z_j^{(s)} = \bar{V}_j + e_j^{(s)} + \sigma \cdot \zeta(x^{(s)}), \forall j = 1, \ldots, J$.

5. Sample $\epsilon_{j^\star}$ for the purchased option. The draw needs to satisfy both the continue to search condition and part of the choice condition $u_{j^\star} > 0$.

   a. Draw $\epsilon_{j^\star}^{(s)}$ from its distribution with upper bound $\bar{b} = -\bar{V}_{j^\star} - e_{j^\star}^{(s)} - \omega_{j^\star} + z_J^{(s)}$ and lower bound $\underline{b} = \max(\max_{j' \in (1,\ldots,J)}$ $z_{j'}^{(s)}, 0) - \bar{V}_{j^\star} - e_{j^\star}^{(s)} - \omega_{j^\star}$. We get $u_{j^\star}^{(s)} = \bar{V}_{j^\star} + e_{j^\star}^{(s)} + \omega_{j^\star} + \epsilon_{j^\star}^{(s)}$.

6. We then evaluate the likelihood that the choice condition $u_{j^\star}$ is larger than the utility of the other searched option. The probability that ϵ_j's, $j \neq j^\star$ satisfy the inequalities can be directly evaluated.

   a. Note that $P(\epsilon_j < \bar{b}_j, \forall j \in (1, \ldots, J) \setminus j^\star) = \prod_{j \in (1,\ldots,J) \setminus j^\star} \Phi(\bar{b}_j)$, where the upper bound is $\bar{b}_j = -\bar{V}_j - e_j^{(s)} - \omega_j + u_{j^\star}^{(s)}$.

7. Finally, the likelihood can be approximated by the average of the probabilities using a set of draws:

$$L^s = \frac{1}{S} \sum_{s=1}^{S} \left\{ \left(1 - \Phi^e\left(\max\left(\max_{j' \notin (1,\ldots,J)}\right.\right.\right.\right.$$

$$\left.\left.\left. \left(-\bar{V}_J + \bar{V}_{j'} + e_{j'}^{(s)}\right), -\bar{V}_J - \sigma \cdot \zeta(x^{(s)})\right)\right)\right) \cdot$$

$$\left(\Phi\left(-\bar{V}_{j^\star} - e_{j^\star}^{(s)} - \omega_{j^\star} + z_J^{(s)}\right) - \Phi\left(\max\left(\max_{j' \notin (1,\ldots,J)} z_{j'}^{(s)}, 0\right)\right.\right.$$

$$\left.\left. - \bar{V}_{j^\star} - e_{j^\star}^{(s)} - \omega_{j^\star}\right)\right) \cdot$$

$$\prod_{j=1}^{J-1} \left(1 - \Phi^e\left(\bar{V}_{j+1} + e_{j+1}^{(s)} - \bar{V}_j\right)\right)$$

$$\left. \cdot \prod_{j \in (1,\ldots,J) \setminus j^\star} \Phi\left(-\bar{V}_j - e_j^{(s)} - \omega_j + u_{j^\star}^{(s)}\right)\right\}.$$

Endnotes

[1] Taobao has since run some retargeting activities that are based on consumers' past behavior. As far as we know, it has not utilized the behavioral factors related to search, which is what we focus on in this study. The search variables, as we show in this paper, can improve the efficiency of retargeting policies.

[2] We use an exchange rate of US$1 = ¥6.2 during March 2013, our sample period.

[3] The data contain consumers' search and purchase activities for a period of 36 days. To guarantee that we have the complete search sequence of each consumer, we exclude individuals who searched in the first week as well as the last week, as their search might have started before or continue after the data observation period. By doing so, we distinguish consumers subject to retargeting from those who finish browsing and come back to purchase later. We define a search process starting from the time when a consumer clicks the first seller link and finishing at the time she clicks the last. It may involve multiple days. In the selected sample, however, only 2.5% of individuals search over multiple days.

[4] Similar to many online retail platforms, Taobao uses a proprietary algorithm to decide the ranking of each seller link. The algorithm relies on many seller-specific factors that do not frequently fluctuate.

[5] See De Los Santos et al. (2012) and Honka and Chintagunta (2016), who test between sequential and simultaneous search models when consumers conduct price search.

[6] This is consistent with the ranking algorithm typically on online retail platforms. The ranking of a seller link is positively correlated with its past sales and conversion; see, for example, https://service.alibaba.com/hc/supplier/detail/20139266.htm (accessed January 8, 2020; in Chinese).

[7] The term $K - 1$ is used instead of K to correct for sample bias.

[8] Empirically, we stop the iteration when $|\zeta_{i,k+1} - \zeta_{ik}| < e^{-10}$.

[9] In practice, researchers can calibrate scaling factors using simulation studies and apply the calibrated scaling factors in estimation (see Ursu 2018).

[10] Drawing from ϕ^e with bounds $(\underline{b}, \bar{b})$ is straightforward. One can draw x from the standard normal distribution ϕ with bounds $(\frac{\underline{b}}{\sigma_e}, \frac{\bar{b}}{\sigma_e})$, and $x \cdot \sigma_e$ follows the distribution ϕ^e with bounds $(\underline{b}, \bar{b})$. For the CDF, it is easy to show that $\Phi^e(\bar{b}) = \Phi(\frac{\bar{b}}{\sigma^e})$.

[11] When $J = 1$, the continue to search condition does not apply. The lower bound $\underline{b} = \max_{j' \notin (1,\dots,J)}(-\bar{V}_J + \bar{V}_{j'} + e_{j'}^{(s)})$.

[12] This assumption will be invalid if ξ_j is correlated with the price or the seller star. In this case, the estimated difference between the high and low groups will pick up the systematic difference in the average ξ's between the two groups. To correctly infer the coefficients for price or for seller stars, however, is not the main focus of this paper. We will discuss the impact on our counterfactual results if the estimated price coefficient is biased in the next section.

[13] We find that the likelihood is maximized when $\sigma_e = 0.5$.

[14] Note that although ω's do not affect *which option* to click, they do affect whether consumers will continue to search. Recall that consumers have rational expectations of the distribution that ω's are drawn from (see Section 3.1 for details). Thus the benefits of searching is smaller if all the ω's are close to 0 than if the ω's are varied in size.

[15] Note that the mean of ξ's and ω's are both normalized to 0. Let sellers 1 and 2 be the high-price sellers, and let sellers 1 and 4 be the high-star sellers. The mean of ξ's among the high-price sellers is normalized to 0: $\xi_1 + \xi_2 = 0$. The mean of ξ's among the high-star-level sellers is also normalized to 0: $\xi_1 + \xi_4 = 0$. Therefore, after normalization, we estimate ξ_1 and ξ_3, and we get $\xi_2 = -\xi_1$, $\xi_4 = -\xi_1$, and $\xi_5 = \xi_1 - \xi_3$. For the ω's, we estimate $\omega_1, \dots, \omega_4$, and we get $\omega_5 = 1 - \sum_1^4 \omega_j$.

[16] This is the typical referral fee charged by Amazon.com for independent sellers. Taobao charges a commission rate ranging from 0.3% to 5% depending on the product category (see Alibaba's 2015 annual report at https://otp.investis.com/clients/us/alibaba/SEC/sec-show.aspx?Type=html&FilingId=11407357&CIK=0001577552&Index=10000,

accessed 8/29, 2020; in Chinese). The results that we later present will remain qualitatively the same if we use these lower rates.

[17] This assumption implies that (1) the consumer remembers all of the seller's information she has searched and (2) the consumer has not purchased from other places, so her need still exists. The assumption is reasonable if the marketing actions are taken shortly after the consumer stops searching. It also implies that (3) the consumer does not update her expectation of the seller's value after she is retargeted.

[18] This coupon value will maximize Taobao's expected profit, assuming that the retail platform provides the service of sending out the coupon, and sellers bid for the service via second price auctions. Details about the auction mechanism will be discussed in Section 6.1. Taobao's profit is equal to the expected revenue from sending coupons plus the increase of referral fees. We use a grid search method to find the optimal coupon value.

[19] Because retargeting does not have any impact on sellers' original profit, it should not cause dissatisfaction among sellers who are not willing to pay Taobao for delivering coupons or recommending sellers. To test the robustness of this conclusion, we rerun the counterfactuals with the price coefficient for ¥100 imposed to be either −0.045 or −0.019 using the estimates from Ursu (2018). Although the magnitude is different, results are similar to Tables 7 and 8. Our conclusions remain the same. Detailed results are available from the authors upon request.

[20] When $J = 1$, the condition to search condition does not apply. The lower bound is $\underline{b} = \max_{j' \notin (1,\dots,J)}(-\bar{V}_J + \bar{V}_{j'} + e_{j'}^{(s)})$.

References

Acquisti A, Varian HR (2005) Conditioning prices on purchase history. *Marketing Sci.* 24(3):367–381.

Bentley T, Chan T, Cosguner K, Park Y-H (2018) Ranking as an imperfect signal in sponsored search advertising and its empirical implications. Working paper, University of Texas at Austin, Austin.

Berke A, Fulton G, Vaccarello L (2014) *The Retargeting Playbook: How To Turn Web-Window Shoppers into Customers* (John Wiley & Sons, Hoboken, NJ).

Bleier A, Eisenbeiss M (2015) Personalized online advertising effectiveness: The interplay of what, when, and where. *Marketing Sci.* 34(5):669–688.

Chan TY, Park Y-H (2015) Consumer search activities and the value of ad positions in sponsored search advertising. *Marketing Sci.* 34(4):606–623.

Chen Y, Yao S (2016) Sequential search with refinement: Model and application with click-stream data. *Management Sci.* 63(12): 4345–4365.

Chen Y, Narasimhan C, Zhang ZJ (2001) Individual marketing with imperfect targetability. *Marketing Sci.* 20(1):23–41.

Choi M, Dai AY, Kim K (2018) Consumer search and price competition. *Econometrica* 86(4):1257–1281.

Cosguner K, Chan TY, Seetharaman PB (2016) Dynamic pricing in a distribution channel in the presence of switching costs. *Management Sci.* 64(3):1212–1229.

Cristal G (2015) *Ad Serving Technology: Understand the Marketing Revelation That Commercialized the Internet* (Gregory Cristal, London).

De Los Santos B, Hortaçsu A, Wildenbeest MR (2012) Testing models of consumer search using data on web browsing and purchasing behavior. *Amer. Econom. Rev.* 102(6):2955–2980.

Elrod T, Winer RS (1982) An empirical evaluation of aggregation approaches for developing market segments. *J. Marketing* 46(4): 65–74.

Fudenberg D, Tirole J (2000) Customer poaching and brand switching. *RAND J. Econom.* 31(4):634–657.

Fudenberg D, Villas-Boas JM (2006) Behavior-based price discrimination and customer recognition. Hendershott T, ed. *Handbook on Economics and Information Systems*, vol. 1 (Emerald Group Publishing, Bingley, UK), 377–436.

Geweke J, Keane M, Runkle D (1994) Alternative computational approaches to inference in the multinomial probit model. *Rev. Econom. Statist.* 76(4):609–632.

Goldfarb A, Tucker CE (2011a) Online display advertising: Targeting and obtrusiveness. *Marketing Sci.* 30(3):389–404.

Goldfarb A, Tucker CE (2011b) Privacy regulation and online advertising. *Management Sci.* 57(1):57–71.

Hajivassiliou VA, Ruud PA (1994) Classical estimation methods for LDV models using simulation. Engle RF, McFadden DL, eds. *Handbook of Econometrics*, vol. 4 (Elsevier, Amsterdam), 2383–2441.

Honka E (2014) Quantifying search and switching costs in the US auto insurance industry. *RAND J. Econom.* 45(4):847–884.

Honka E, Chintagunta P (2016) Simultaneous or sequential? Search strategies in the U.S. auto insurance industry. *Marketing Sci.* 36(1):21–42.

Keane MP (1994) A computationally practical simulation estimator for panel data. *Econometrica* 62(1):95–116.

Kim JB, Albuquerque P, Bronnenberg BJ (2010) Online demand under limited consumer search. *Marketing Sci.* 29(6):1001–1023.

Koulayev S (2014) Search for differentiated products: Identification and estimation. *RAND J. Econom.* 45(3):553–575.

Lambrecht A, Tucker C (2013) When does retargeting work? Information specificity in online advertising. *J. Marketing Res.* 50(5): 561–576.

Pazgal A, Soberman D (2008) Behavior-based discrimination: Is it a winning play, and if so, when? *Marketing Sci.* 27(6): 977–994.

Rossi PE, McCulloch RE, Allenby GM (1996) The value of purchase history data in target marketing. *Marketing Sci.* 15(4): 321–340.

Sahni NS (2016) Advertising spillovers: Evidence from online field experiments and implications for returns on advertising. *J. Marketing Res.* 53(4):459–478.

Sahni NS, Narayanan S, Kalyanam K (2019) An experimental investigation of the effects of retargeted advertising: The role of frequency and timing. *J. Marketing Res.* 56(3):401–418.

Shaffer G, Zhang ZJ (1995) Competitive coupon targeting. *Marketing Sci.* 14(4):395–416.

Shin J, Sudhir K (2010) A customer management dilemma: When is it profitable to reward one's own customers? *Marketing Sci.* 29(4):671–689.

Stigler GJ (1961) The economics of information. *J. Political Econom.* 69(3):213–225.

Ursu RM (2018) The power of rankings: Quantifying the effect of rankings on online consumer search and purchase decisions. *Marketing Sci.* 37(4):530–552.

Villas-Boas JM (1999) Dynamic competition with customer recognition. *RAND J. Econom.* 30(4):604–631.

Villas-Boas JM (2004) Consumer learning, brand loyalty, and competition. *Marketing Sci.* 23(1):134–145.

Weitzman ML (1979) Optimal search for the best alternative. *Econometrica* 47(3):641–654.